



STIC Search Report

Biotech-Chem Library

File Copy
09/20/04, 4:39
updated

STIC Database Tracking Number: 140315

TO: David Lamberston
Location: rem/2b79/2c70
Art Unit: 1636
Tuesday, December 14, 2004

Case Serial Number: 09/864486

From: Barb O'Bryen
Location: Biotech-Chem Library
Remsen 1A69
Phone: 571-272-2518

barbara.obryen@uspto.gov

Search Notes

O'Bryen, Barbara

From: Lambertson, David
Sent: Monday, December 13, 2004 8:35 AM
To: O'Bryen, Barbara
Cc: Lambertson, David
Subject: Search Request

Search Request

Examiner's Name:	David Lambertson
Examiner #:	79514
Art Unit:	1636
Room #:	02B79 Remsen
Mailbox room#:	02C70 Remsen
Phone:	(571) 272-0771
Results Format:	paper

Serial #:09/864,486

Please Search:

Nucleic Acid and Protein databases for:

SEQ ID No: 1 and 2

Including:

1. Default Search.
2. Interference Search.

Thanks,
Dave.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: December 13, 2004, 16:30:37 ; Search time 63 Seconds

(without alignments)
4099.763 Million cell updates/sec

Title:

Perfect score: 599

Sequence: 1 gtaatagaggtctctaagta.....catcttttagccctggaac 360

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 4004546

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlh
-Q=/cgn2_1/usPTO_spool/US09864486/runat_13122004_100258_1951/app_query.fasta_1.910
-DB=A Geneseq 23Sep04 -OPMT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09864486 @CGN_1_156 @runat_13122004_100258_1951 -NCPU=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A Geneseq 23Sep04:*

1: Geneseqp1980s: *
2: Geneseqp1990s: *
3: Geneseqp2000s: *
4: Geneseqp2001s: *
5: Geneseqp2002s: *
6: Geneseqp2003as: *
7: Geneseqp2003bs: *
8: Geneseqp2004s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	110	18.6	58	ABR62697	ABR62697 Yeast GCN
C 2	110	18.6	281	AAW02621	AAW02621 Yeast GCN
C 3	110	18.6	281	AAU08953	AAU08953 Yeast GCN
C 4	110	18.6	281	ABG94229	ABG94229 Yeast GCN
C 5	110	18.6	281	ABG93091	ABG93091 S. cerevi
C 6	110	18.6	281	ABG80541	ABG80541 Yeast gen
C 7	110	18.6	281	ADB67084	ADB67084 General c
C 8	110	18.6	281	ADL99355	ADL99355 Nanostruc
C 9	109	18.4	323	ABG93326	ABG93326 C. albica
C 10	102	17.2	49	ABU09710	ABU09710 Saccharom

C 11	98	16.6	266	2	AA91295	AA91295 Drosophil
C 12	92	15.5	36	8	ADF94459	ADF94459 Structure
C 13	91	15.4	281	2	AAW13955	AAW13955 LZ varian
C 14	90.5	15.3	36	3	AAAY88676	AAAY88676 Core poly
C 15	90.5	15.3	36	4	AAAB77031	AAAB77031 Core poly
C 16	90.5	15.3	36	4	ABBO1499	ABBO1499 Viral cor
C 17	90.5	15.3	36	4	ABB00035	ABB00035 Viral DP1
C 18	90.5	15.3	36	4	AAU12584	AAU12584 DP178-lik
C 19	90.5	15.3	36	5	ADE01519	ADE01519 Hybrid po
C 20	89	15.0	35	4	AAG66864	AAG66864 Leucine z
C 21	89	15.0	35	6	ABG75995	ABG75995 Leucine z
C 22	88	15.0	35	6	AAE32714	AAE32714 Dimerisat
C 23	88	14.9	281	2	AAW02622	AAW02622 Yeast GCN
C 24	87	14.7	33	3	AAAY88745	AAAY88745 Core poly
C 25	87	14.7	33	4	AAE07425	AAE07425 Yeast PDA
C 26	87	14.7	33	4	ABBO1569	ABBO1569 Viral cor
C 27	87	14.7	33	4	ABBO0104	ABBO0104 Viral DP1
C 28	87	14.7	33	4	AAU12653	AAU12653 DPI78-lik
C 29	87	14.7	33	4	AAU12653	AAU12653 DPI78-lik
C 30	87	14.7	33	5	AAE15760	AAE15760 PDA-3Hb p
C 31	87	14.7	33	5	AAU76769	AAU76769 PDA-3Hb p
C 32	87	14.7	33	5	ABG75547	ABG75547 Optimised
C 33	87	14.7	33	5	ADE01589	ADE01589 Hybrid po
C 34	87	14.7	33	8	ADF94454	ADF94454 Structure
C 35	87	14.7	35	8	ADMA6699	ADMA6699 GNC4 leuc
C 36	87	14.7	36	2	AAW95616	AAW95616 Humanized
C 37	87	14.7	36	2	AAW30633	AAW30633 Recombina
C 38	87	14.7	36	4	AAAB6778	AAAB6778 Leucine z
C 39	87	14.7	36	8	ADQ94457	ADQ94457 Structure
C 40	87	14.7	43	8	ADQ28739	ADQ28739 Linked 80
C 41	87	14.7	44	2	AAO6684	AAO6684 Sequence
C 42	87	14.7	44	8	ADO39716	ADO39716 Fab'2 ant
C 43	87	14.7	45	2	AAW40588	AAW40588 Human Fab
C 44	87	14.7	45	2	AAW92434	AAW92434 Plasmid p
C 45	87	14.7	45	2	AAAY29455	AAAY29455 GCN4 leuc

ALIGNMENTS

RESULT 1

ABR62697

ID ABR62697 standard; protein; 58 AA.

XX ABR62697;

XX 06-NOV-2003 (first entry)

XX Yeast GCN4 basic domain leucine zipper domain.

XX Yeast; GCN4; biliverdin reductase; enzyme; haem oxygenase-1;

XX antiasthmatic; antiinflammatory; antiulcer; vasotropic; hypertensive;

XX antibacterial; gene therapy; antisense therapy; leucine zipper.

XX Saccharomyces cerevisiae.

XX Key Location/Qualifiers

FT Domain 18..23

FT Domain /label= Basic

FT Domain 30..58

FT Misc-difference 30 /label= Leucine_zipper

FT Misc-difference 37 /note= "L1 residue of leucine zipper"

FT Misc-difference 44 /note= "L2 residue of leucine zipper"

FT Misc-difference 51 /note= "L3 residue of leucine zipper"

FT Misc-difference 58 /note= "L4 residue of leucine zipper"

FT Misc-difference 58 /note= "L5 residue of leucine zipper"

XX WO2003055981-A2.

XX

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: December 13, 2004, 18:14:18 ; Search time 55.8 Seconds
(without alignments)
4608.753 Million cell updates/sec

Title: US-09-864-486A-1
Perfect score: 599
Sequence: 1 gaataggagtctctaagta.....catcttctagcccttggaaac 360

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1585576 seqs, 357178320 residues

Total number of hits satisfying chosen parameters: 3171152

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -MODEL=frame+ n2p.model -DEV=xlh
-O=/cgn2_1/USPTO.spool/US09864486/runat_13122004_100303_2053/app.query.fasta_1.910
-DB=Published Applications AA -QWTF=fastan -SUFFIX=rapb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -THR SCORE=pct -THR MAX=100
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -END=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09864486 @CNG 1.168 @runat_13122004_100303_2053
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOC=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-Fgapop=6 -Fgapext=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pcp:
2: /cgn2_6/ptodata/1/pubpaa/PT_NEW_PUB.pcp:
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pcp:
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pcp:
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pcp:
6: /cgn2_6/ptodata/1/pubpaa/PTUS_PUBCOMB.pcp:
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pcp:
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pcp:
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pcp:
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pcp:
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pcp:
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pcp:
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pcp:
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pcp:
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pcp:
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pcp:
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pcp:
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pcp:
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pcp:
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pcp:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result Query

No.	Score	Match	Length	DB	ID	Description
C 1	110	18.6	281	9	US-09-801-368-120	Sequence 120, App
C 2	110	18.6	281	10	US-09-848-616-154	Sequence 154, App
C 3	110	18.6	281	14	US-10-050-902-154	Sequence 154, App
C 4	110	18.6	281	14	US-10-050-998-154	Sequence 154, App
C 5	110	18.6	281	14	US-10-080-608A-15	Sequence 15, Appl
C 6	110	18.6	281	14	US-10-370-685-104	Sequence 104, App
C 7	110	18.6	281	16	US-10-451-467A-140	Sequence 140, App
C 8	109	18.4	323	16	US-10-451-467A-610	Sequence 610, App
C 9	102	17.2	49	15	US-10-622-064-12	Sequence 12, Appl
C 10	92	15.5	36	14	US-10-283-403-6	Sequence 6, Appl
C 11	90.5	15.3	36	14	US-10-351-641-26	Sequence 26, Appl
C 12	89	15.0	35	13	US-10-057-505-17	Sequence 17, Appl
C 13	87	14.7	33	14	US-10-351-641-96	Sequence 96, Appl
C 14	87	14.7	33	14	US-10-283-403-1	Sequence 1, Appl
C 15	87	14.7	33	17	US-10-851-564-4	Sequence 4, Appl
C 16	87	14.7	36	9	US-09-940-166A-3	Sequence 3, Appl
C 17	87	14.7	36	14	US-10-283-403-4	Sequence 4, Appl
C 18	87	14.7	36	16	US-10-762-967-3	Sequence 3, Appl
C 19	87	14.7	43	14	US-10-347-960-26	Sequence 26, Appl
C 20	87	14.7	45	10	US-09-726-258-57	Sequence 57, Appl
C 21	87	14.7	48	14	US-10-247-960-29	Sequence 29, Appl
C 22	87	14.7	55	14	US-10-247-960-5	Sequence 5, Appl
C 23	87	14.7	55	17	US-10-730-776-14	Sequence 14, Appl
C 24	87	14.7	60	14	US-10-247-960-6	Sequence 6, Appl
C 25	87	14.7	60	17	US-10-730-776-15	Sequence 15, Appl
C 26	87	14.7	298	10	US-09-726-258-60	Sequence 60, Appl
C 27	87	14.7	300	9	US-09-940-166A-7	Sequence 7, Appl
C 28	87	14.7	300	14	US-10-227-694-2	Sequence 2, Appl
C 29	87	14.7	300	16	US-10-762-967-7	Sequence 7, Appl
C 30	83	14.0	52	15	US-10-344-620-2	Sequence 2, Appl
C 31	83	14.0	191	15	US-10-344-620-6	Sequence 6, Appl
C 32	82	13.9	32	10	US-09-259-658-5	Sequence 5, Appl
C 33	82	13.9	32	10	US-09-491-614-4	Sequence 4, Appl
C 34	82	13.9	32	14	US-10-103-597A-4	Sequence 4, Appl
C 35	82	13.9	32	14	US-10-161-205-3	Sequence 3, Appl
C 36	82	13.9	32	14	US-10-101-001-3	Sequence 3, Appl
C 37	82	13.9	32	14	US-10-188-444-4	Sequence 4, Appl
C 38	82	13.9	33	9	US-09-796-202-6	Sequence 6, Appl
C 39	82	13.9	33	14	US-10-032-314-6	Sequence 316, App
C 40	82	13.9	33	14	US-10-323-314-6	Sequence 6, Appl
C 41	82	13.9	298	14	US-10-273-180-4	Sequence 4, Appl
C 42	81	13.7	33	14	US-10-338-083-18	Sequence 18, Appl
C 43	81	13.7	33	16	US-10-611-399-18	Sequence 18, Appl
C 44	81	13.7	36	14	US-10-283-403-8	Sequence 8, Appl
C 45	81	13.7	52	15	US-10-395-817-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-09-801-368-120
; Sequence 120, Application US/09801368
; Patent No. US20020128250A1
; GENERAL INFORMATION:
; APPLICANT: Busby, Robert
; APPLICANT: Cali, Brian
; APPLICANT: Hecht, Peter
; APPLICANT: Holtzman, Doug
; APPLICANT: Madden, Kevin
; APPLICANT: Maxon, Mary
; APPLICANT: Milne, Todd
; APPLICANT: No. US20020128250Alman, Thea
; APPLICANT: Royer, John
; APPLICANT: Salama, Sofie
; APPLICANT: Sherman, Amir
; APPLICANT: Silva, Jeff
; APPLICANT: Summers, Eric
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
; FILE REFERENCE: 109272.147
; CURRENT APPLICATION NUMBER: US/09/801,368
; CURRENT FILING DATE: 2001-03-07

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: December 13, 2004, 17:56:18 : Search time 18 Seconds
(without alignments)
3848.671 Million cell updates/sec

Title: US-09-864-486A-1
Perfect score: 599
Sequence: 1 gtaataggagtcttaagta.....catcttcttagcccttgaac 360

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+_n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09864486/runat_13122004_100300_1969/app_query.fasta_1.910
-DB=PIR_79 -QFMT=fastan -SUFFIX=rpr -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=500 -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09864486 @CGN_1_1_35 @runat_13122004_100300_1969 -NCPU=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PIR_79.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	110	18.6	281	1 RGBYA2	amino acid biosynt
c 2	89	15.0	2201	1 GNNYA9	genome polyprotein
c 3	81	13.7	2185	1 GNNYET	genome polyprotein
c 4	81	13.7	2185	1 GNNYB3	genome polyprotein
c 5	75	12.7	1091	2 T13170	diaphanous protein
c 6	70	11.8	582	2 S22195	bps2 protein - Des
c 7	69.5	11.7	2206	1 GNNY4P	genome polyprotein
c 8	69.5	11.7	2206	1 GNNY27	genome polyprotein
c 9	68.5	11.6	2207	1 G38122	genome polyprotein
c 10	68.5	11.6	2209	1 GNNY3P	genome polyprotein
c 11	68.5	11.6	2209	1 GNNY2P	genome polyprotein
c 12	68.5	11.3	401	2 AB0864	probable aminotran
c 13	67.5	11.3	2214	1 A48548	genome polyprotein
c 14	67	11.3	2214	1 A48548	genome polyprotein

15	66	11.0	83	2 AE1913	hypothetical prote
16	66	11.0	664	2 C84747	probable protein k
c 17	66	11.1	1379	2 S37310	protoporphyrin IX
c 18	65	11.0	412	1 KRSHU1	keratin, 48K type
c 19	65	11.0	416	2 A61404	keratin A, type I
c 20	65	11.0	416	2 S60034	keratin H ₁ , type I
c 21	65	11.0	416	2 A46559	keratin, type I, h
c 22	65	10.9	628	2 T05900	phosphoenolpyruvat
c 23	65	10.9	2405	2 T08164	dynein alpha heavy
c 24	64.5	10.9	1140	1 S38908	UV-damaged DNA-bin
c 25	64.5	10.9	1140	1 S38977	UV-damaged DNA-bin
c 26	64.5	10.9	1140	2 JC7152	UV-damaged DNA-bin
c 27	64	10.8	354	2 T06487	probable DNA-bindi
c 28	64	10.7	355	2 F96940	beta-mannanase (im
c 29	64	10.8	357	2 S50821	DNA-binding protei
c 30	64	10.8	593	2 A71532	probable phosphoma
c 31	64	10.7	670	2 S52637	phosphoenolpyruvat
c 32	64	10.8	2182	1 GNNYB1	genome polyprotein
c 33	63.5	10.6	144	2 G69425	hypothetical prote
c 34	63.5	10.7	547	2 F69964	amino acid degrada
c 35	63.5	10.6	714	2 S66699	hypothetical prote
c 36	63	10.6	190	2 B72416	conserved hypothet
c 37	63	10.6	296	1 TVFVJN	transforming prote
c 38	63	10.6	327	2 F71980	hypothetical prote
c 39	63	10.6	392	2 A60777	keratin 2, type I,
c 40	63	10.6	404	2 JS0073	keratin, 47.6k typ
c 41	63	10.6	405	2 T45615	hypothetical prote
c 42	62.5	10.6	269	1 A30208	cross-pathway cont
c 43	62.5	10.4	280	2 T28225	ORF MSV064 hypothe
c 44	62.5	10.6	329	2 F86675	mevalonate kinase
c 45	62.5	10.4	739	2 A46159	interferon-depende

ALIGNMENTS

RESULT 1
RGBYA2
amino acid biosynthesis regulatory protein - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein YEL009C
C;Species: Saccharomyces cerevisiae
C;Date: 27-Nov-1985 #sequence revision 27-Nov-1985 #text_change 09-Jul-2004
C;Accession: A03605; S50450; A03604
R;Hinnebusch, A.G.
Proc. Natl. Acad. Sci. U.S.A. 81, 6442-6446, 1984
A;Title: Evidence for translational regulation of the activator of general amino acid
A;Reference number: A03605; MUID:85038531; PMID:6387704
A;Accession: A03605
A;Molecule type: DNA
A;Residues: 1-281 <HIN>
A;Cross-references: UNIPROT:P03069; EMBL:K02205; NID:gl71581; PIDN:AAA34640.1; PID:gl7
R;Dietrich, F.S.
submitted to the EMBL Data Library, December 1994
A;Description: Saccharomyces cerevisiae chromosome V cosmids 9871, 8199, 9867, 9495 an
A;Reference number: S50428
A;Accession: S50450
A;Molecule type: DNA
A;Residues: 1-281 <DIE>
A;Cross-references: EMBL:U18530; NID:g602367; PIDN:AAB64486.1; PID:g602376; GSPDB:GN00
R;Thiross, G.; Penn, M.D.; Greer, H.
Proc. Natl. Acad. Sci. U.S.A. 81, 5096-5100, 1984
A;Title: 5' untranslated sequences are required for the translational control of a yeast
A;Reference number: A03604; MUID:84298088; PMID:6433345
A;Accession: A03604
A;Molecule type: DNA
A;Residues: 1-238, 'PGVLVRESCKE' <THI>
A;Cross-references: EMBL:K02649; NID:gl71585; PIDN:AAA65521.1; PID:gl71586
C;Genetics:
A;Gene: SGD:GCN4; MIPS:YEL009C
A;Cross-references: SGD:S0000735; MIPS:YEL009C
A;Map position: 5L
C;Function:
A;Description: transcription regulation
A;Note: required under amino acid starvation conditions for increasing the transcripti

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: December 13, 2004, 16:32:22 ; Search time 81.9 Seconds
(without alignments)
5058.239 Million cell updates/sec

Title: US-09-864-486A-1
Perfect score: 599
Sequence: 1 gtaattaggagttcttaagta.....catctttttgacctggaac 360

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 3650362

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlh
-Q=/cgn2.1/USPTO.spool/US09864486/runat.13122004.100259.1961/app.query.fasta_1.910
-DB=uniprot_02 -QFMT=fastan -SUFFIX=rup -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -NATRIX=blosum62 -FRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pcio -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09864486 @CGN 1.1.141 @runat.13122004.100259.1961 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : UniProt 02.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	119	20.1	281	Q6C4D2	O6C4d2 yarrowia li
C 2	118	19.9	325	Q75AC9	Q75ac9 ashbya goss
C 3	118	19.9	325	AAS51909	Aas51909 ashbya go
C 4	112	18.9	314	Q877C4	Q877c4 candida mal
C 5	112	18.9	333	Q6CQU7	Q6cqu7 kluyveromyc
C 6	110	18.6	281	1 GCN4 YEAST	P03069 saccharomyc
C 7	110	18.6	281	Q70D88	Q70d88 saccharomyc
C 8	110	18.6	281	Q70D91	Q70d91 saccharomyc
C 9	110	18.6	281	Q70D96	Q70d96 saccharomyc
C 10	110	18.6	281	Q70D99	Q70d99 saccharomyc
C 11	110	18.6	281	Q70DA0	Q70da0 saccharomyc
C 12	110	18.6	281	CAE52206	Caes52206 saccharomyc
C 13	110	18.6	281	CAE52207	Caes52207 saccharomyc
C 14	110	18.6	281	CAE52208	Caes52208 saccharomyc
C 15	110	18.6	281	CAE52209	Caes52209 saccharomyc
C 16	110	18.6	281	CAE52210	Caes52210 saccharomyc

C 17	110	18.6	281	2	CAE52211	Caes52211 saccharom
C 18	110	18.6	281	2	CAE52212	Caes52212 saccharom
C 19	110	18.6	281	2	CAE52213	Caes52213 saccharom
C 20	110	18.6	281	2	CAE52214	Caes52214 saccharom
C 21	110	18.6	281	2	CAE52215	Caes52215 saccharom
C 22	110	18.6	281	2	CAE52216	Caes52216 saccharom
C 23	110	18.6	281	2	CAE52217	Caes52217 saccharom
C 24	110	18.6	281	2	CAE52218	Caes52218 saccharom
C 25	110	18.6	281	2	CAE52219	Caes52219 saccharom
C 26	110	18.6	281	2	CAE52220	Caes52220 saccharom
C 27	110	18.6	281	2	CAE52221	Caes52221 saccharom
C 28	110	18.6	281	2	CAE52222	Caes52222 saccharom
C 29	110	18.6	281	2	CAE52223	Caes52223 saccharom
C 30	110	18.6	281	2	CAE52224	Caes52224 saccharom
C 31	109	18.4	322	2	Q9UV12	Q9uv12 candida alb
C 32	106	17.9	310	2	Q6FL16	Q6fl16 candida gla
C 33	94	15.9	287	2	Q6BW79	Q6bw79 debaryomyce
C 34	89	15.0	118	2	Q6VD37	Q6vd37 human coxs
C 35	89	15.0	118	2	Q6VD43	Q6vd43 human coxs
C 36	89	15.0	118	2	Q6VD45	Q6vd45 human coxs
C 37	89	15.0	118	2	Q6VD83	Q6vd83 human coxs
C 38	89	15.0	118	2	Q6VD84	Q6vd84 human coxs
C 39	89	15.0	118	2	Q6VD85	Q6vd85 human coxs
C 40	89	15.0	118	2	Q6VD86	Q6vd86 human coxs
C 41	89	15.0	118	2	Q6VDG7	Q6vdg7 human coxs
C 42	89	15.0	118	2	Q6VDG8	Q6vdg8 human coxs
C 43	89	15.0	118	2	Q6VDK6	Q6vdk6 human coxs
C 44	89	15.0	118	2	Q6VDL0	Q6vdl0 human coxs
C 45	89	15.0	118	2	AAQ95293	Aaq95293 human cox

ALIGNMENTS

RESULT 1	Q6C4D2	PRELIMINARY;	PRT;	281 AA.
ID	Q6C4D2	PRELIMINARY;	PRT;	281 AA.
AC	Q6C4D2	PRELIMINARY;	PRT;	281 AA.
DT	01-OCT-2004	(TREMELREL. 28, Created)		
DT	01-OCT-2004	(TREMELREL. 28, Last sequence update)		
DT	01-OCT-2004	(TREMELREL. 28, Last annotation update)		
DE	Chromosome E of strain CLIB99 of Yarrowia lipolytica.			
GN	ORFNames=YALI0E27742g;			
OS	Yarrowia lipolytica (Candida lipolytica).			
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;			
OC	Saccharomycetales; Dipodascaceae; Yarrowia.			
OX	NCBI_TaxID=4952;			
RN	[1]_TaxID=4952;			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=CLIB99;			
RG	GENOLEVURES;			
RA	Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,			
RA	Lafontaine I., de Montigny J., Marck C., Neuvéglise C., Talla E., V.,			
RA	Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,			
RA	Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,			
RA	Boisrame A., Boyer J., Cattolico L., Confanioli F., de Daruvar A.,			
RA	Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,			
RA	Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,			
RA	Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,			
RA	Nicaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,			
RA	Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,			
RA	Swennene D., Tekai F., Wesolowski-Louvel M., Westhof E., Wirth B.,			
RA	Zenou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,			
RA	Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,			
RT	Wincker P., Souciet J.L.;			
RT	"Genome evolution in yeasts.;"			
RL	Nature 430:35-44(2004).			
RL	[2]			
RC	STRAIN=CLIB99;			
RP	Genoscope;			
RA	Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.			
RL	EMBL; CR382131; CAG80083.1;			
SEQUENCE	281 AA; 30378 MW; 16BF8B2D3AF86C96 CRC64;			

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 13, 2004, 14:03:27 ; Search time 2078.4 Seconds
(without alignments)
8191.056 Million cell updates/sec

Title: US-09-864-486A-1
Perfect score: 360
Sequence: 1 gtaataggagttcctaagta.....catcttcttagccctggaac 360

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :					GenEmbl.*				
					1: gb_ba.*				
					2: gb_htg.*				
					3: gb_in.*				
					4: gb_om.*				
					5: gb_ov.*				
					6: gb_pat.*				
					7: gb_ph.*				
					8: gb_pl.*				
					9: gb_pr.*				
					10: gb_ro.*				
					11: gb_scs.*				
					12: gb_sy.*				
					13: gb_un.*				
					14: gb_vi.*				
					Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.				
					SUMMARIES				
Result No.	Score	Query %	Match	Length	DB	ID	Description		
1	360	100.0	360	6	BD175189	DNA fragm			
2	360	100.0	360	6	AX323247	Sequence			
3	360	100.0	2696	8	PPA272040	Pichia pa			
C 4	62.8	17.4	1824	6	AR064394	Sequence			
C 5	62.8	17.4	1824	6	I28708	Sequence 3			
C 6	62.8	17.4	1824	6	I89362	Sequence 3			
C 7	62.8	17.4	1824	8	YSCGCN4	Yeast (S. ce			
C 8	62.8	17.4	2089	8	YSCGCN4B	Yeast (S. ce			
C 9	62.8	17.4	62643	8	SCB9871	Yeast (S. ce			
C 10	62.8	17.4	110000	2	AC141230_2	Continuation [3 of			
C 11	54	15.0	1580	8	AF205716	Candida a			
C 12	53.4	14.8	972	6	AX537008	Sequence			
C 13	53	14.7	846	8	AJ585686	Saccharom			
C 14	53	14.7	846	8	AJ585687	Saccharom			
C 15	53	14.7	846	8	AJ585688	Saccharom			
C 16	53	14.7	846	8	AJ585689	Saccharom			
C 17	53	14.7	846	8	AJ585690	Saccharom			
C 18	53	14.7	846	8	AJ585691	Saccharom			
C 19	53	14.7	846	8	AJ585692	Saccharom			

C 20	53	14.7	846	8	AJ585693	Saccharom
C 21	53	14.7	846	8	AJ585694	Saccharom
C 22	53	14.7	846	8	AJ585695	Saccharom
C 23	53	14.7	846	8	AJ585696	Saccharom
C 24	53	14.7	846	8	AJ585697	Saccharom
C 25	53	14.7	846	8	AJ585698	Saccharom
C 26	53	14.7	846	8	AJ585699	Saccharom
C 27	53	14.7	846	8	AJ585700	Saccharom
C 28	53	14.7	846	8	AJ585701	Saccharom
C 29	53	14.7	846	8	AJ585702	Saccharom
C 30	53	14.7	846	8	AJ585703	Saccharom
C 31	53	14.7	846	8	AJ585704	Saccharom
C 32	53	14.7	1346	6	AX536538	Sequence
C 33	52.2	14.5	2800	8	AB063247	AB063247 Candida m
C 34	51.6	14.3	110000	8	CR380958_02	Continuation (3 of
C 35	47.4	13.2	110000	8	CR382134_10	Continuation (11 o
C 36	47	13.1	7104	6	AX281477_10	AX281477 Sequence
C 37	47	13.1	7104	6	AX348827	AX348827 Sequence
C 38	45.2	12.6	95202	10	AL928807	AJ928807 Mouse DNA
C 39	44.8	12.4	110000	8	CR382124_12	Continuation (13 o
C 40	44.2	12.3	798	6	AR473525	AR473525 Sequence
C 41	44.2	12.3	110000	8	CR382131_32	Continuation (33 o
C 42	44.2	12.3	132728	2	AC099411_1	AC099411 Felis cat
C 43	44.2	12.3	220251	2	AC150374	AC150374 Colobus g
C 44	43.2	12.0	6763	6	AX251241	AX251241 Sequence
C 45	43.2	12.0	6763	6	AX281459	AX281459 Sequence

ALIGNMENTS

RESULT 1
BD175189
LOCUS DNA fragment of a mechiru nutrition nature picha pathtorisu yeast
DEFINITION ICL gene.
ACCESSION BD175189
VERSION BD175189.1 GI:29120883
KEYWORDS JP 2002253235-A/2.
SOURCE Pichia pastoris
ORGANISM Pichia pastoris
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Pichia.
REFERENCE 1 (bases 1 to 360)
AUTHORS Dias,J.M., Prado,I.V. and Leon,N.C.
TITLE DNA fragment of a mechiru nutrition nature picha pathtorisu yeast
JOURNAL ICL gene
COMMENT Patent: JP 2002253235-A 2 10-SEP-2002;
CENTRO DE INGENIERIA GENETICA Y BIOTECNOLOGIA
OS Pichia pastoris (Yeast)
PN JP 2002253235-A/2
PD 10-SEP-2002
PF 28-MAY-2001 JP 2001159321
PR 26-MAY-2000 CU 2000-122
PI JAVIER MENEZES DIAS,IRIS VALDEZ PRADO,NELSON CABRERA LEON PC
C12N15/09,C12N15/00
CC Secuencia que contiene la region 3' no codificante del gen CC
FH Key ICL. Location/Qualifiers
FT terminator Location/Qualifiers
source 1..360
/organism="Pichia pastoris"
/mol_type="genomic DNA"
/db_xref="taxon:4922"

ORIGIN
Query Match 100.0%; Score 360; DB 6; Length 360;
Best Local Similarity 100.0%; Pred. No. 1.3e-76;
Matches 360; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTATAGAGTCTCCTAGTAGTTAAGTAATAATTGACTTGAGCTATTATAGATTGTGTGT 60
|||||

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 13, 2004, 14:02:37 ; Search time 339 Seconds

(without alignments)
5574.606 Million cell updates/sec

Title: US-09-864-486A-1

Perfect score: 360

Sequence: 1 gtaataggaggttcctaagta.....catcttcttagccctggaac 360

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_23Sep04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	358.4	99.6	360	ABK13231	Abk13231 P. pastor
2	62.8	17.4	1824	AAT32832	Aat32832 Yeast GCN
3	53.4	14.8	972	ABQ76592	Abq76592 C. albica
4	53	14.7	1346	ABQ76357	Abq76357 S. cerevi
5	52.2	14.5	2684	ADP57624	Adf57624 Yeast GCN
6	47	13.1	7104	ABL34588	AbL34588 Human met
7	47	13.1	7104	ABL70395	AbL70395 Chemicall
8	44.2	12.3	798	AAT14025	Aat14025 Drosophil
9	43.2	12.0	6763	AAS46487	Aas46487 Tumour su
10	43.2	12.0	6763	ABL33554	AbL33554 Human imm
11	43.2	12.0	6763	ABL34570	AbL34570 Human met
12	43.2	12.0	6763	ABN80196	Abn80196 Human che
13	43	11.9	175	AAAL46025	Aal46025 Peptide d
14	43	11.9	175	AAAL46026	Aal46026 pGADDA424
15	43	11.9	586	AAAL46021	Aal46021 NLS and t
16	42	11.7	156	AAAL46020	Aal46020 DNA bindi
17	41.8	11.6	8771	ABL33825	AbL33825 Human imm
18	41.6	11.6	12763	ABL32302	AbL32302 Human imm
19	41.4	11.5	504	ABQ73929	Abq73929 Chimeric
20	41.4	11.5	525	ABQ73935	Abq73935 TBD-ci-bz
21	41.4	11.5	542	ABQ73934	Abq73934 TBD-ci-bz

C 22	41.2	11.4	5020	8	ABZ10105	Abz10105 Haematopo
C 23	40.8	11.3	8718	6	ABL33273	AbL33273 Human imm
C 24	40.4	11.2	6196	6	ABL33864	AbL33864 Human imm
C 25	40.2	11.2	606	5	ABV49913	Abv49913 Human pro
C 26	40	11.1	897	12	ADI23637	Adi23637 Yeast GCN
C 27	40	11.1	5489	6	ABL34120	AbL34120 Human imm
C 28	39.4	10.9	10133	6	ABL32458	AbL32458 Human imm
C 29	39.2	10.9	966	6	ABL87946	AbL87946 DNA polym
C 30	39.2	10.9	966	6	ABL87945	AbL87945 DNA polym
C 31	38.8	10.8	6912	6	ABK28371	AbK28371 DNA trans
C 32	38.6	10.7	16633	6	ABN79984	Abn79984 Human che
C 33	38.6	10.7	18434	6	ABL34007	AbL34007 Human imm
C 34	38.6	10.7	53155	6	ABN87364	Abn87364 Human lip
C 35	38.4	10.7	6316	4	AAS46351	Aas46351 Tumour su
C 36	38.4	10.7	6316	6	ABK31260	AbK31260 Signal tr
C 37	38.4	10.7	6316	6	ABL70215	AbL70215 Chemicall
C 38	38.4	10.7	6321	6	AAS61170	Aas61170 Human gen
C 39	38.4	10.7	6321	6	ABQ67063	AbQ67063 Human ang
C 40	38.2	10.6	9770	6	ABL32032	AbL32032 Human imm
C 41	38.2	10.6	83391	6	ABQ67093	AbQ67093 Human ang
C 42	38	10.6	4993	6	ABL34103	AbL34103 Human imm
C 43	38	10.6	5020	8	ABZ09959	Abz09959 Haematopo
C 44	38	10.6	5234	6	ABL32945	AbL32945 Human imm
C 45	38	10.6	6250	6	ABN80215	Abn80215 Human che

ALIGNMENTS

RESULT 1

ABK13231
ID ABK13231 standard; DNA; 360 BP.

AC ABK13231;

DT 23-APR-2002 (first entry)

DE P. pastoris isocitrate lyase gene 3' UTR.

KW Isocitrate lyase; ICL; 3' UTR; ds; yeast; heterologous gene expression.

XX Pichia pastoris.

PN BP1162366-A2.

PD 12-DEC-2001.

PF 28-MAY-2001; 2001EP-00202009.

PR 26-MAY-2000; 2000CU-00000122.

XX (INGG-) CENT ING GENETICA & BIOTECNOLOGIA.

XX Menendez Diaz J, Valdes Prado I, Cabrera Leon N;

DR WPI; 2002-084415/12.

PT New recombinant DNA fragment for protein production, contains an isocitrate lyase encoding gene from Pichia pastoris, which regulates heterologous gene expression in yeasts, when operably linked to the fragment.

XX Claim 1; Page 8; 18pp; English.

PS The invention relates to a recombinant DNA fragment (I) comprising a sequence (S) of 684bp (the 5' region of the Pichia pastoris isocitrate lyase, ICL, gene including the promoter) or 360bp (the 3' UTR including a terminator), given in the specification, where (I) belongs to an isocitrate lyase encoding gene (ICL) from Pichia pastoris, which is able to regulate heterologous gene expression in yeasts, when it is operably linked to (I), and including regulatory elements necessary for heterologous gene expression. The DNA fragment is useful for expressing heterologous proteins in Pichia pastoris, for expressing a foreign gene.

GenCore version 5.1.6
Copyright (C) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 13, 2004, 14:11:07 ; Search time 67.8 Seconds
(without alignments)
3774.100 Million cell updates/sec

Title: US-09-864-486A-1
Perfect score: 360
Sequence: 1 gtaataggagtcttaagta.....catcttcttagcccttggaac 360

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PTCUS_COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	62.8	17.4	1824	1 US-08-347-792-3	Sequence 3, Appli
C 2	62.8	17.4	1824	1 US-08-431-357-3	Sequence 3, Appli
C 3	62.8	17.4	1824	2 US-08-697-221-5	Sequence 5, Appli
C 4	62.8	17.4	1824	5 PCT-US95-15353-3	Sequence 3, Appli
C 5	44.2	14.4	945	4 US-09-248-796A-4712	Sequence 4712, Ap
C 6	44.2	12.3	798	2 US-08-319-866-7	Sequence 7, Appli
C 7	44.2	12.3	798	4 US-08-809-917-7	Sequence 7, Appli
C 8	37.2	10.3	1824	1 US-08-347-792-5	Sequence 5, Appli
C 9	37.2	10.3	1824	1 US-08-431-357-5	Sequence 5, Appli
C 10	37.2	10.3	1824	5 PCT-US95-15353-5	Sequence 5, Appli
C 11	37.2	10.3	2523	4 US-09-620-312D-290	Sequence 290, App
C 12	36	10.0	525	4 US-08-956-171E-539	Sequence 539, App
C 13	36	10.0	525	4 US-08-781-986A-539	Sequence 539, App
C 14	36	10.0	9919	3 US-08-880-179-1	Sequence 1, Appli
C 15	35.8	9.9	10144	4 US-10-204-708-94	Sequence 94, Appli
C 16	35.4	9.8	5666	4 US-10-204-708-29	Sequence 29, Appli
C 17	35.2	9.8	399	4 US-09-621-976-8976	Sequence 8976, Ap
C 18	35.2	9.8	3300	3 US-08-913-842-4	Sequence 4, Appli
C 19	34.8	9.7	151	2 US-08-256-790-3	Sequence 3, Appli
C 20	34.8	9.7	1141	4 US-09-806-708B-22	Sequence 22, Appli
C 21	34.8	9.7	7664	4 US-10-204-708-83	Sequence 83, Appli
C 22	34.8	9.7	10467	4 US-10-204-708-1	Sequence 1, Appli
C 23	34.4	9.6	8961	4 US-10-204-708-79	Sequence 79, Appli
C 24	34.2	9.5	1064	3 US-09-149-922-3	Sequence 3, Appli
C 25	33.8	9.4	1055	4 US-09-806-708B-23	Sequence 23, Appli
C 26	33.8	9.4	6583	4 US-10-204-708-26	Sequence 26, Appli
C 27	33.6	9.3	5152	4 US-10-204-708-48	Sequence 48, Appli

C 28	33.6	9.3	9347	4 US-10-204-708-35	Sequence 35, Appli
C 29	33.6	9.3	11050	4 US-10-204-708-86	Sequence 86, Appli
C 30	33.6	9.3	84495	3 US-09-797-906-3	Sequence 3, Appli
C 31	33.4	9.3	10144	4 US-10-204-708-93	Sequence 93, Appli
C 32	33.4	9.3	11049	4 US-10-204-708-23	Sequence 23, Appli
C 33	33.2	9.2	138	4 US-09-376-463-25	Sequence 25, Appli
C 34	33.2	9.2	138	4 US-09-376-463-27	Sequence 27, Appli
C 35	33	9.2	153	4 US-09-376-463-28	Sequence 28, Appli
C 36	33	9.2	153	4 US-09-376-463-30	Sequence 30, Appli
C 37	33	9.2	1030	4 US-09-270-767-11977	Sequence 11977, A
C 38	33	9.2	1410	3 US-08-936-165A-190	Sequence 190, App
C 39	33	9.2	1994	4 US-09-899-463A-31	Sequence 31, Appli
C 40	33	9.2	6583	4 US-10-204-708-45	Sequence 45, Appli
C 41	33	9.2	19233	4 US-10-204-708-25	Sequence 25, Appli
C 42	32.8	9.1	855	4 US-09-248-796A-2979	Sequence 2979, Ap
C 43	32.8	9.1	1215	4 US-09-248-796A-3573	Sequence 3573, Ap
C 44	32.8	9.1	1548	4 US-09-248-796A-5684	Sequence 5684, Ap
C 45	32.8	9.1	193303	4 US-09-497-855A-37	Sequence 37, Appli

ALIGNMENTS

RESULT 1
US-08-347-792-3/c
; Sequence 3, Application US/08347792
; Patent No. 5573925
; GENERAL INFORMATION:
; APPLICANT: Halazonetis, Thanos D.
; TITLE OF INVENTION: p53 Proteins With Altered
; TITLE OF INVENTION: Tetramerization Domains
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr., PO Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/347,792
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: WSTS8USA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9206
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1824 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 778..1623
US-08-347-792-3

Query Match 17.4%; Score 62.8; DB 1; Length 1824;
Best Local Similarity 63.0%; Pred. No. 2.6e-08;
Matches 97; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 207 AAATCAATATAAAAAAACTTACATCATGTGTTTCACTAACTCTTTAAACGC 266

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 13, 2004, 15:27:22 ; Search time 354 Seconds
(without alignments)
5606.970 Million cell updates/sec

Title: US-09-864-486A-1
Perfect score: 360
Sequence: 1 gtaataggaggtcctaagta.....catcttttagcccttgaac 360

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4085731 seqs, 2756760397 residues

Total number of hits satisfying chosen parameters: 8171462

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
- 20: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query	ID	Description
1	360	100.0	360	11	US-09-864-486A-1
2	240	66.7	240	11	US-09-864-486A-2
3	53.4	14.8	972	17	US-10-451-467A-609
4	53	14.7	846	9	US-09-801-368-119
5	53	14.7	1346	17	US-10-451-467A-139
6	47	13.1	7104	15	US-10-240-485-141
7	43.2	12.0	6763	15	US-10-311-455-1527
8	43.2	12.0	6763	15	US-10-240-485-123
9	43.2	12.0	6763	16	US-10-221-714A-209
10	43	11.9	175	16	US-10-344-620-10
11	43	11.9	175	16	US-10-344-620-11
12	43	11.9	586	16	US-10-344-620-5

C 13	42	11.7	156	16	US-10-344-620-4
C 14	41.8	11.6	8771	15	US-10-311-455-1798
C 15	41.6	11.6	12763	15	US-10-311-455-275
C 16	41.4	11.5	504	14	US-10-032-827A-12
C 17	41.4	11.5	525	14	US-10-032-827A-18
C 18	41.4	11.5	542	14	US-10-032-827A-17
C 19	41.2	11.4	5020	18	US-10-473-126-245
C 20	40.8	11.3	8718	15	US-10-473-126-245
C 21	40.4	11.2	6196	15	US-10-311-455-1837
C 22	40	11.1	897	15	US-10-311-455-1837
C 23	40	11.1	5489	15	US-10-311-455-2093
C 24	39.4	10.9	10133	15	US-10-311-455-431
C 25	39.4	10.9	3673778	15	US-10-312-841-2
C 26	39.2	10.9	966	10	US-09-906-179A-16
C 27	39.2	10.9	966	10	US-09-906-179A-17
C 28	38.8	10.8	288	18	US-10-425-115-3622
C 29	38.8	10.8	6912	15	US-10-240-453-245
C 30	38.6	10.7	388	17	US-10-437-963-1934
C 31	38.6	10.7	18434	15	US-10-311-455-1980
C 32	38.6	10.7	55155	9	US-09-735-933-3
C 33	38.4	10.7	6316	16	US-10-221-613-129
C 34	38.4	10.7	6316	16	US-10-221-714A-73
C 35	38.4	10.7	6321	17	US-10-433-793-93
C 36	38.2	10.6	9770	15	US-10-311-455-5
C 37	38.2	10.6	83391	17	US-10-433-793-123
C 38	38	10.6	689	13	US-10-027-632-185044
C 39	38	10.6	689	15	US-10-027-632-185044
C 40	38	10.6	4993	15	US-10-311-455-2076
C 41	38	10.6	5020	18	US-10-473-126-99
C 42	38	10.6	5234	15	US-10-311-455-918
C 43	38	10.6	7591	15	US-10-074-024-841
C 44	37.8	10.5	6123	15	US-10-311-455-794
C 45	37.8	10.5	3673778	15	US-10-312-841-1

ALIGNMENTS

RESULT 1
US-09-864-486A-1
; Sequence 1, Application US/09864486A
; Publication No. US20040106783A1
; GENERAL INFORMATION:
; APPLICANT: Menendez Diaz, Javier
; APPLICANT: Valdes Prado, Chris
; APPLICANT: Cabrera Leon, Nelson
; TITLE OF INVENTION: DNA Fragments of the Methylotrophic Pichia Pastoris Yeast iCL 9
; FILE REFERENCE: 976-9
; CURRENT APPLICATION NUMBER: US/09/864,486A
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 360
; TYPE: DNA
; ORGANISM: Pichia pastoris
; US-09-864-486A-1

Query Match	100.0%;	Score 360;	DB 11;	Length 360;
Best Local Similarity	100.0%;	Pred. No. 4.5e-83;		
Matches 360;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	GTAATAGGAGTTCCTAAGTAGTAAAGATAATTCAGTTGAGGTATTTATAGATTGTGTGT	60	
Db	1	GTAATAGGAGTTCCTAAGTAGTAAAGATAATTCAGTTGAGGTATTTATAGATTGTGTGT	60	
Qy	61	AGGTAATATCTATGTCGTCCATTCCTTACCTTGGTGGGTGACGGGCGGTGAATAATC	120	
Db	61	AGGTAATATCTATGTCGTCCATTCCTTACCTTGGTGGGTGACGGGCGGTGAATAATC	120	
Qy	121	AGTTGGCATCAAGACTTTTACACCTTGTCCACAGAGGGTCCGCTCTACTGATTACTACA	180	
Db	121	AGTTGGCATCAAGACTTTTACACCTTGTCCACAGAGGGTCCGCTCTACTGATTACTACA	180	

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	61.4	17.1	323	7	T36997	T36997 EST102040 S
2	61.4	17.1	363	7	T36351	T36351 EST101276 S
3	61	16.9	263	7	T37110	T37110 EST102177 S
4	60.8	16.3	370	7	T36326	T36326 EST101246 S
5	59.8	16.6	320	7	T38676	T38676 EST104204 S
6	57	15.8	210	7	T38672	T38672 EST104198 S
7	48.8	13.6	987	9	CNS06PHH	AL409371 T3 end of
8	46	12.8	476	1	AI399548	AI399548 NCSP6C277
9	44.8	12.4	440	2	AW870991	AW870991 ra53d02.y
10	44.8	12.4	542	2	AW827656	AW827656 ra45c07.y
11	44.4	12.3	904	9	CNS06GNCH	AL406599 T3 end of
12	43.8	12.2	430	1	AU267015	AU267015 AU267015
13	43.8	12.2	459	6	C93400	C93400 C93400 Dict
14	43.8	12.2	487	1	AU267016	AU267016 AU267016
15	43.4	12.1	611	9	CE651766	CE651766 tigr-gss-
16	42.2	11.7	937	9	CL054187	CL054187 Ch216-78N
17	42	11.7	438	8	BZ286848	BZ286848 SALK_0201
18	41.8	11.6	605	6	CF308474	CF308474 ABF--02-F
19	41.8	11.6	870	9	CNS06UZZ	AL416517 T7 end of
20	41.6	11.6	297	9	CNS03NEJ	AL251956 Tetraodon
21	41.2	11.4	388	8	BZ382828	BZ382828 SALK_1189
22	41.2	11.4	610	2	BB770903	BB770903 BB770903
23	40.8	11.3	1101	9	CNS000L9Y	AL068053 Drosophil
24	40.8	11.3	1101	9	CNS0002B7	AL097453 Drosophil

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: December 13, 2004, 16:30:37 ; Search time 42 Seconds

(without alignments)

4099.763 Million cell updates/sec

Title: US-09-864-486A-2

Perfect score: 402

Sequence: 1 gtaataggagtcttaagta.....aaaaaaacttaacatcact 240

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 4004546

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-O=/cgm2.1/USPTO.spool/US09864486/runat.13122004.100258.1951/app.query.fasta_1.910
-DB=A_Geneseq.23Sep04 -QWTF=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09864486 @CGN 1.1.156 @runat.13122004.100258.1951 -NCPU=6 -ICPU=3
-NO_MMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A_Geneseq.23Sep04:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	81	20.0	2185	2 AAR12141	Aar12141 Enterovir
C 2	67.5	16.6	2206	2 AAR22210	Aar22210 True type
C 3	66.5	16.4	2209	1 AAP20037	Aap20037 Sequence
C 4	66	16.3	157	6 ADA34042	Ada34042 Acinetoba
5	65	16.2	564	4 ABB66319	Abb66319 Drosophil
6	65	16.2	576	3 AAG39442	Aag39442 Arabidops
7	65	16.2	670	3 AAG39441	Aag39441 Arabidops
8	64.5	16.0	54	4 ABG08201	Abg08201 Novel hum
C 9	64.5	15.9	1140	2 AAR72386	Aar72386 XAP-1, pa
C 10	64.5	15.9	1140	4 AAU69747	Aau69747 Human xer

C 11	64.5	15.9	1140	5 ABG31550	Abg31550 Human dam
C 12	64.5	15.9	1140	5 ABB78348	Abb78348 Amino aci
C 13	64.5	15.9	1140	8 ADJ66572	Adj66572 UV-damage
C 14	64	15.8	593	2 AAY36891	Aay36891 C. tracho
15	64	15.9	618	4 AAG67089	Aag67089 Shiitake
16	64	15.9	627	2 AAW62553	Aaw62553 Protein o
17	64	15.9	627	2 AAW62552	Aaw62552 Shiitake
18	64	15.9	653	2 AAW14264	Aaw14264 Z. japoni
C 19	63	15.5	405	3 AAG09410	Aag09410 Arabidops
C 20	63	15.5	411	3 AAG09409	Aag09409 Arabidops
21	63	15.7	658	2 AAW14263	Aaw14263 S. anglic
C 22	62.5	15.4	329	5 ABB53713	Abb53713 Lactococc
C 23	62.5	15.4	1568	8 ADO42011	Ado42011 Human cel
24	62	15.4	471	6 ABU28669	Abu28669 Protein e
25	62	15.4	1689	4 AEG16013	Aeg16013 Novel hum
26	62	15.4	2417	4 AEG28727	Aeg28727 Novel hum
27	61.5	15.3	403	4 AAU29734	Aau29734 Novel hum
C 28	61	15.0	408	8 ADK16139	Adk16139 Nancarcha
29	61	15.2	649	3 AAG36263	Aag36263 Arabidops
30	61	15.2	671	3 AAG36262	Aag36262 Arabidops
31	61	15.2	671	5 ABB93280	Abb93280 Herbicida
32	61	15.2	671	6 ABR39578	Ab39578 A. thalia
33	61	15.2	671	7 ADF28852	Adf28852 A. thalia
C 34	60.5	14.9	208	4 AAM16746	Aam16746 Peptide #
C 35	60.5	14.9	208	4 ABB35732	Abb35732 Peptide #
C 36	60.5	14.9	208	4 AAM29234	Aam29234 Peptide #
C 37	60.5	14.9	208	4 ABB30567	Abb30567 Peptide #
C 38	60.5	14.9	208	4 ABB21158	Abb21158 Protein #
C 39	60.5	14.9	208	4 AAM68923	Aam68923 Human bon
C 40	60.5	14.9	208	4 AAM56541	Aam56541 Human bra
C 41	60.5	14.9	208	4 ABG50588	Abg50588 Human liv
C 42	60.5	14.9	208	4 AAM04463	Aam04463 Peptide #
C 43	60.5	14.9	208	5 ABG38504	Abg38504 Human pep
44	60	14.9	420	3 AAY69561	Aay69561 Trametes
45	60	14.9	420	3 AAB20518	Aab20518 Trametes

ALIGNMENTS

RESULT 1

AAR12141
ID AAR12141 standard; protein; 2185 AA.
XX

AC AAR12141;

DT 25-MAR-2003 (revised)

DT 05-AUG-1991 (first entry)

XX Enteroviral polypeptide.

DE Enteroviral polypeptide.

XX Enteroviruses; monoclonal antibodies; myocarditis; meningo-

XX encephalitis; pancreatitis; post viral fatigue.

OS Enterovirus sp.

XX DE3939200-A.

XX 29-MAY-1991.

XX 27-NOV-1989; 89DE-03939200.

XX 27-NOV-1989; 89DE-03939200.

XX (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.

XX Kandolf R;

XX WPI; 1991-165150/23.

XX N-ESDB; AAQ11816.

XX New enteroviral polypeptide for raising group specific antibodies - for

XX detecting any type of enterovirus in blood or serum, also new DNA

XX encoding it.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: December 13, 2004, 17:52:37 ; Search time 11.6 Seconds
(without alignments)

2744.193 Million cell updates/sec

Title: US-09-864-486A-2

Perfect score: 402

Sequence: 1 gtaatggaggttcctaagta.....aaaaaaacttaacatcact 240

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 956278

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -MODEL-frame+ n2p.model -DEV=xlh

-Q=/cgn2_1/USFTO.spool/US09864486/runat_13122004_100300_1979/app_query.fasta_1.910

-DB=Issued Patents AA -OFMT=fastan -SUFFIX=rai -MINMATCH=0.1 -LOOPCL=0

-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi

-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15

-MODE=LOCAL -OUTPMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000

-USER=US09864486 @CNG 1.1.35 @runat_13122004_100300_1979 -NCPU=6 -ICPU=3

-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSFBLOCK=100 -LONGLOG

-DEV_TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6

-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents AA.*

1: /cgn2_6/prodata/1/iaa/5A_COMB.pep.*

2: /cgn2_6/prodata/1/iaa/5B_COMB.pep.*

3: /cgn2_6/prodata/1/iaa/6A_COMB.pep.*

4: /cgn2_6/prodata/1/iaa/6B_COMB.pep.*

5: /cgn2_6/prodata/1/iaa/PCTUS_COMB.pep.*

6: /cgn2_6/prodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	67.5	16.6	2206	1	US-07-852-260-2
C 2	67.5	16.6	2206	2	US-08-461-503-2
C 3	67.5	16.6	2206	3	US-08-465-250-2
C 4	66	16.3	157	4	US-09-328-352-5329
C 5	65	16.2	504	4	US-09-270-767-46764
C 6	65	16.2	542	4	US-09-270-767-42841
C 7	64.5	15.9	1140	3	US-09-651-656-21
C 8	64.5	15.9	1140	3	US-09-650-855-21
C 9	61	15.0	91	4	US-09-248-796A-27715
C 10	60	14.9	410	4	US-09-684-855-143
C 11	60	14.9	420	4	US-09-684-855-120
C 12	60	14.9	420	4	US-09-488-265B-19

13	60	14.9	443	3	US-08-993-359-30	Sequence 30, Appl
14	60	14.9	443	4	US-09-273-871A-5	Sequence 5, Appli
15	60	14.9	443	4	US-09-482-558A-30	Sequence 30, Appli
16	60	14.9	443	4	US-10-083-452-5	Sequence 5, Appli
C 17	59.5	14.7	828	4	US-09-540-236-3107	Sequence 3107, Ap
C 18	59	14.7	828	4	US-09-489-039A-8129	Sequence 8129, Ap
C 19	59	14.5	547	4	US-09-555-000-2	Sequence 2, Appli
C 20	58.5	14.4	859	4	US-09-978-522-3	Sequence 3, Appli
C 21	58	14.3	328	4	US-09-248-796A-20603	Sequence 20603, A
C 22	58	14.3	442	4	US-09-684-855-137	Sequence 137, App
C 23	57.5	14.3	715	4	US-09-328-352-5229	Sequence 5229, Ap
C 24	57	14.2	191	4	US-09-248-796A-16718	Sequence 16718, A
C 25	57	14.0	216	4	US-09-270-767-45610	Sequence 45610, A
C 26	57	14.0	224	4	US-09-248-796A-20864	Sequence 20864, A
C 27	56.5	14.1	376	3	US-09-387-418A-13	Sequence 13, Appli
C 28	56.5	14.1	582	4	US-09-430-808A-3	Sequence 3, Appli
C 29	56.5	13.9	643	4	US-09-270-767-40779	Sequence 40779, A
C 30	56.5	13.9	643	4	US-09-270-767-55995	Sequence 55995, A
C 31	56.5	14.1	712	1	US-08-369-796-6	Sequence 6, Appli
C 32	56.5	14.1	712	2	US-08-852-091-6	Sequence 6, Appli
C 33	56.5	14.1	712	2	US-08-820-754-6	Sequence 6, Appli
C 34	56.5	14.1	712	3	US-08-956-652-6	Sequence 6, Appli
C 35	56.5	14.1	712	3	US-08-956-652-6	Sequence 6, Appli
C 36	56.5	14.1	712	3	US-08-948-547-6	Sequence 6, Appli
C 37	56.5	14.1	712	3	US-08-956-653A-6	Sequence 6, Appli
C 38	56.5	14.1	712	4	US-08-212-185-6	Sequence 6, Appli
C 39	56.5	14.1	712	4	US-09-430-808A-2	Sequence 2, Appli
C 40	56.5	14.1	712	5	PCT-US95-17025-6	Sequence 6, Appli
C 41	56.5	14.1	729	4	US-09-917-254-97	Sequence 97, Appl
C 42	56.5	14.1	740	1	US-08-276-099A-12	Sequence 12, Appl
C 43	56.5	14.1	740	1	US-08-781-890-12	Sequence 12, Appl
C 44	56.5	14.1	750	1	US-08-369-796-4	Sequence 4, Appli
C 45	56.5	14.1	750	2	US-08-852-091-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-07-852-260-2
; Sequence 2, Application US/07852260
; Patent No. 5525715
; GENERAL INFORMATION:
; APPLICANT: Racanelli, Vincent
; APPLICANT: Tatem, Joanne M.
; APPLICANT: Weeks-Levy, Carolyn L.
; TITLE OF INVENTION: METHODS FOR PRODUCING RNA VIRUSES FROM
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Cooper & Dunham
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/852,260
; FILING DATE: 19920619
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 36607-B-PCT-US
; TELEPHONE: (212) 977-9550
; TELEFAX: (212) 664-0525
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 2:

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model
Run on: December 13, 2004, 18:14:18 ; Search time 37.2 Seconds
(without alignments)
4608.753 Million cell updates/sec

Title: US-09-864-486A-2
Perfect score: 402
Sequence: 1 gaaataggagttcctaagta.....aaaaaaacttaacatcact 240

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1585576 seqs, 357178320 residues
Total number of hits satisfying chosen parameters: 3171152

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09864486/runat_13122004_100303_2053/app_query.fasta_1.910
-DB=published Applications AA -QFW=fastan -SUFFIX=rapb -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USPR=US09864486 @CGN 1 1 168 @runat_13122004_100303_2053
-NCPU=6 -ICPU=3 -NO MWAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOC=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pcp:
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pcp:
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pcp:
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pcp:
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pcp:
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pcp:
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pcp:
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pcp:
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pcp:
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pcp:
11: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pcp:
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pcp:
13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pcp:
14: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pcp:
15: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pcp:
16: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pcp:
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pcp:
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pcp:
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pcp:
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pcp:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result * Query

No.	Score	Match	Length	DB	ID	Description
1	70	17.4	170	16	US-10-767-701-49593	Sequence 49593, A
2	70	17.4	386	17	US-10-425-115-194870	Sequence 194870,
3	69.5	17.1	157	15	US-10-424-599-236702	Sequence 236702,
4	69	17.0	187	15	US-10-425-114-44299	Sequence 44299, A
5	67.5	16.6	148	15	US-10-424-599-150125	Sequence 150125,
6	67.5	16.6	154	17	US-10-425-115-218584	Sequence 218584,
7	67	16.5	89	15	US-10-424-599-242199	Sequence 242199,
8	67	16.7	663	16	US-10-437-963-188657	Sequence 188657,
9	64.5	15.9	89	17	US-10-425-115-297092	Sequence 297092,
10	64.5	15.9	1140	16	US-10-468-406-8	Sequence 8, Appli
11	64	15.8	325	15	US-10-424-599-254766	Sequence 254766,
12	63	15.5	413	15	US-10-425-114-39092	Sequence 39092, A
13	62.5	15.4	112	17	US-10-425-115-222873	Sequence 222873,
14	62.5	15.4	329	14	US-10-369-493-18334	Sequence 18334, A
15	62	15.4	104	15	US-10-424-599-169421	Sequence 169421,
16	62	15.3	231	15	US-10-424-599-273714	Sequence 273714,
17	62	15.4	471	15	US-10-282-122A-56593	Sequence 56593, A
18	62	15.4	636	16	US-10-437-963-184530	Sequence 184530,
19	61	15.0	108	15	US-10-424-599-179712	Sequence 179712,
20	61	15.2	671	14	US-10-217-939-14	Sequence 14, Appl
21	60.5	14.9	120	15	US-10-425-114-43581	Sequence 43581, A
22	60.5	14.9	208	9	US-09-864-761-36456	Sequence 36456, A
23	60	14.9	410	14	US-10-442-538-143	Sequence 143, App
24	60	14.9	420	14	US-10-442-538-120	Sequence 120, App
25	60	14.9	443	10	US-09-999-214-30	Sequence 30, Appl
26	60	14.9	443	13	US-10-083-452-5	Sequence 5, Appli
27	60	14.9	443	16	US-10-734-510-5	Sequence 5, Appli
28	60	14.9	962	16	US-10-437-963-187589	Sequence 187589,
29	59.5	14.7	73	15	US-10-424-599-185823	Sequence 185823,
30	59.5	14.7	85	15	US-10-424-599-188894	Sequence 188894,
31	59.5	14.7	120	15	US-10-424-599-212424	Sequence 212424,
32	59.5	14.7	737	15	US-10-425-114-57283	Sequence 57283, A
33	59.5	14.7	747	15	US-10-282-122A-63122	Sequence 63122, A
34	59	14.5	174	15	US-10-406-686A-53	Sequence 53, Appl
35	59	14.7	218	14	US-10-369-493-2768	Sequence 2768, Ap
36	59	14.7	472	15	US-10-282-122A-59548	Sequence 59548, A
37	58.5	14.6	207	15	US-10-264-049-2701	Sequence 2701, Ap
38	58.5	14.6	225	15	US-10-424-599-177258	Sequence 177258,
39	58.5	14.4	859	10	US-09-978-522-3	Sequence 3, Appli
40	58	14.3	63	16	US-10-437-963-142126	Sequence 142126,
41	58	14.4	391	15	US-10-282-122A-55899	Sequence 55899, A
42	58	14.4	400	14	US-10-369-493-1129	Sequence 1129, Ap
43	58	14.4	442	14	US-10-442-538-137	Sequence 137, App
44	58	14.4	450	15	US-10-282-122A-43729	Sequence 43729, A
45	58	14.4	450	15	US-10-282-122A-76233	Sequence 76233, A

ALIGNMENTS

RESULT 1
US-10-767-701-49593
; Sequence 49593, Application US/10767701
; Publication NO. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 49593
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3476-061-P1-K1-B10.pcp
US-10-767-701-49593

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: December 13, 2004, 17:56:18 ; Search time 12 seconds

(without alignments)
3848.671 Million cell updates/sec

Title: US-09-864-486A-2

Perfect score: 402

Sequence: 1 gtaataggagtcttaagta.....aaaaaaacttaacatcact 240

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283416 segs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlh
-Q/cgna2_1/USPRO_spool/US09864486/runat_13122004_100300_1969/app query.fasta_1.910
-DB=PIR_79 -QFMT=fastan -SUFFIX=rpr -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09864486 -CGN 1_1_35 @runat_13122004_100300_1969 -NCPU=6 -ICPU=3
-NO_WMAP -LARGESQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

PIR_79:.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	88	21.7	2201	1	genome polyprotein
C 2	81	20.0	2185	1	genome polyprotein
C 3	81	20.0	2185	1	genome polyprotein
C 4	67.5	16.8	401	2	probable aminotran
C 5	67.5	16.6	2206	1	genome polyprotein
C 6	67.5	16.6	2206	2	genome polyprotein
C 7	66.5	16.4	2206	1	genome polyprotein
C 8	66.5	16.4	2207	1	genome polyprotein
C 9	66.5	16.4	2209	1	genome polyprotein
C 10	66.5	16.4	2209	1	genome polyprotein
C 11	66	16.3	1379	2	genome polyprotein
C 12	65	16.2	628	2	phosphoenolpyruv
C 13	64.5	15.9	1140	1	UV-damaged DNA-bin
C 14	64.5	15.9	1140	1	UV-damaged DNA-bin

C 15	64.5	15.9	1140	2	JC7152	UV-damaged DNA-bin
C 16	64	15.9	355	2	F96940	beta-mannanase (im
C 17	64	15.8	593	2	A71532	probable phosphona
C 18	64	15.9	670	2	S52637	phosphoenolpyruvat
C 19	64	15.8	2182	1	GNNYB1	genome polyprotein
C 20	63	15.5	405	2	T45615	hypothetical prote
C 21	62.5	15.4	329	2	F86675	mevalonate kinase
C 22	62.5	15.4	2214	1	A48548	genome polyprotein
C 23	62	15.4	471	2	D64973	yegp protein - Esc
C 24	62	15.4	471	2	F90988	probable heat choc
C 25	62	15.4	471	2	H85833	probable heat choc
C 26	62	15.4	634	2	A32241	lactose transport
C 27	62	15.3	2185	1	JQ2021	genome polyprotein
C 28	61	15.0	292	2	AG2225	3-hydroxyacid dehy
C 29	61	15.2	671	2	T06034	phosphoenolpyruvat
C 30	61	15.0	2185	1	GNNYSV	genome polyprotein
C 31	61	15.0	2185	1	GNNYSH	genome polyprotein
C 32	60.5	14.9	1611	2	T38236	hypothetical prote
C 33	60.5	14.9	2207	2	S09553	genome polyprotein
C 34	60	14.8	298	2	B71685	hypothetical prote
C 35	59.5	14.8	406	2	A10292	probable selenocys
C 36	59.5	14.7	2206	1	GNNY21	genome polyprotein
C 37	59	14.7	218	2	S76407	phosphoribosylanth
C 38	59	14.5	547	2	F69964	amino acid degrada
C 39	58	14.3	76	2	T42153	protein kinase hom
C 40	58	14.3	396	2	D97133	hypothetical prote
C 41	58	14.3	400	1	F69142	probable hexosyltr
C 42	58	14.4	450	2	AB0771	conserved hypothet
C 43	58	14.3	5762	2	A41819	proline-rich pepti
C 44	57.5	14.3	483	2	A25896	beta-adrenergic re
C 45	57.5	14.2	576	2	S12792	protein-tyrosine k

ALIGNMENTS

RESULT 1

GNNYA9
Genome polyprotein - coxsackievirus A9 (strain Griggs)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core p
polymerase (EC 2.7.7.48)
C:Species: coxsackievirus A9
C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 09-Jul-2004
C:Accession: JQ0523
F:Chang, K.H.; Auvinen, P.; Hyypia, T.; Stanway, G.
J. Gen. Virol. 70, 3269-3280, 1989
A:Title: The nucleotide sequence of coxsackievirus A9: implications for receptor bindi
A:Reference number: JQ0523; MUID:90111704; PMID:2558158
A:Accession: JQ0523
A:Molecule type: mRNA
A:Residues: 1-2201 <CHA>
A:Cross-references: UNIPROT:P21404
C:Superfamily: poliovirus genome polyprotein
C:Keywords: coat protein; core protein; genome-linked protein; nucleotidyltransferase;
F:1-69/Product: coat protein 1A #status predicted <VP4>
F:70-330/Product: coat protein 1B #status predicted <VP2>
F:331-568/Product: coat protein 1C #status predicted <VP3>
F:569-870/Product: coat protein 1D #status predicted <VP1>
F:871-1017/Product: core protein 2A #status predicted <2AP>
F:1018-1116/Product: core protein 2B #status predicted <2BP>
F:1117-1145/Product: core protein 2C #status predicted <2CP>
F:1446-1534/Product: protein 3A #status predicted <P3A>
F:1535-1556/Product: genome-linked protein VPg #status predicted <VPG>
F:1557-1739/Product: proteinase #status predicted <PTS>
F:1740-2201/Product: RNA-directed RNA polymerase #status predicted <RPS>
F:1537/Binding site: phosphoryl-RNA (Tyr) (covalent) #status predicted

Alignment Scores:

Pred. No.: 0.00222 Length: 2201
Score: 88.00 Matches: 20
Percent Similarity: 54.69% Conservative: 15
Best Local Similarity: 31.25% Mismatches: 23
Query Match: 21.67% Indels: 6
DB: 1 Gaps: 2

GenCore version 5.1.6

OM nucleic - protein search, using frame plus n2p model

Title: US-09-864-486A-2

Scoring table: BLOSUM62

Searched: 1825181 seqs. 575374646 residues

Minimum DB seq length: 0

Post-processing: Minimum Match 0%

Command line parameters:

Database : Uniprot 02: *

Pred. No. is the number of results predicted by chance to have a

SUMMARY

Result No.	Score	Query ϵ		DB	ID	Description
		Match	Length			
C 1	88	21.7	118	2	Q6VD37	human coxsA
C 2	88	21.7	118	2	Q6VD43	human coxsA
C 3	88	21.7	118	2	Q6VD45	human coxsA
C 4	88	21.7	118	2	Q6VD83	human coxsA
C 5	88	21.7	118	2	Q6VD84	human coxsA
C 6	88	21.7	118	2	Q6VD85	human coxsA
C 7	88	21.7	118	2	Q6VD86	human coxsA
C 8	88	21.7	118	2	Q6VDF5	human coxsA
C 9	88	21.7	118	2	Q6VDG7	human coxsA
C 10	88	21.7	118	2	Q6VDG8	human coxsA
C 11	88	21.7	118	2	Q6VDK1	human coxsA
C 12	88	21.7	118	2	Q6VDK6	human coxsA
C 13	88	21.7	118	2	Q6VDL0	human coxsA
C 14	88	21.7	118	2	AAQ95293	human cox
C 15	88	21.7	118	2	AAQ95297	human cox
C 16	88	21.7	118	2	AAQ95302	human cox

C 17	88	21.7	118	2	AAQ95335	Aaq95335	human	cox
C 18	88	21.7	118	2	AAQ95336	Aaq95336	human	cox
C 19	88	21.7	118	2	AAQ95348	Aaq95348	human	cox
C 20	88	21.7	118	2	AAQ95417	Aaq95417	human	cox
C 21	88	21.7	118	2	AAQ95418	Aaq95418	human	cox
C 22	88	21.7	118	2	AAQ95419	Aaq95419	human	cox
C 23	88	21.7	118	2	AAQ95420	Aaq95420	human	cox
C 24	88	21.7	118	2	AAQ95458	Aaq95458	human	cox
C 25	88	21.7	118	2	AAQ95460	Aaq95460	human	cox
C 26	88	21.7	118	2	AAQ95466	Aaq95466	human	cox
C 27	88	21.7	184	2	Q75P21	Q75p21	human	coxsa
C 28	88	21.7	184	2	BAD12600	Bad12600	human	cox
C 29	88	21.7	252	2	Q75P22	Q75p22	human	coxsa
C 30	88	21.7	252	2	Q75P23	Q75p23	human	coxsa
C 31	88	21.7	252	2	Q75P24	Q75p24	human	coxsa
C 32	88	21.7	252	2	BAD12597	Bad12597	human	cox
C 33	88	21.7	252	2	BAD12598	Bad12598	human	cox
C 34	88	21.7	252	2	BAD12599	Bad12599	human	cox
C 35	88	21.7	252	1	POLG_CXA9	P21404	c. genome	po
C 36	87	21.4	118	2	Q5VD22	Q5vd22	human	coxsa
C 37	87	21.4	118	2	AAQ95481	Aaq95481	human	cox
C 38	82	20.2	2185	2	Q9E7C2	Q9e7c2	human	coxsa
C 39	81	20.0	310	2	Q9PY72	Q9py72	human	coxsa
C 40	81	20.0	851	2	Q9PYE3	Q9pye3	human	coxsa
C 41	81	20.0	936	2	Q9E7C1	Q9e7c1	human	coxsa
C 42	81	20.0	1028	2	Q86331	Q86331	human	coxsa
C 43	81	20.0	2185	1	POLG_CXB3N	P03313	c. genome	po
C 44	81	20.0	2185	1	POLG_CXB3W	Q66282	c. genome	po
C 45	81	20.0	2185	2	O66338	O66338	human	coxsa

ALIGNMENTS

[illegible]

Alignment Scores:		
Pred. No.:	0.0128	118
Score:	88.00	20
Percent Similarity:	54.6%	15
Best Local Similarity:	31.2%	23
Query Match:	21.6%	2
DB:	2	2
	Length:	Matches:
	Conservative:	Mismatches:
	Indels:	Gaps:

US-09-864-486A-2 (1-240) x O6VP37 (1-118)

Qy 233 TTAAGTTTTTTTTTTATATTGATTTTGAATTGAAAATTTTTATCCAAAGTCGTTTGTA 174

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 13, 2004, 14:03:27 ; Search time 1385.6 Seconds
(without alignments)
8191.056 Million cell updates/sec

Title: US-09-864-486A-2
Perfect score: 240
Sequence: 1 gtaataggagttcctaagta.....aaaaaaaaacttaacatcact 240

Scoring table: IDENTITY_NUC
Gapop 10_0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues
Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_srs.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	240	100.0	360	6	BD175189 DNA fragm
2	240	100.0	360	6	AX323247 Sequence
3	240	100.0	2696	8	PA272040 Pichia pa
4	44	18.3	93202	10	AL928807 Mouse DNA
5	41	17.1	136268	10	AC068496
6	41	17.1	213482	10	AC073883
7	40.8	17.0	132728	2	AC099411
8	40.6	16.9	198973	2	CR628332
9	39.6	16.5	244962	2	AC098272
10	39.6	16.5	251257	2	AC095972
11	39.2	16.3	128277	2	AC099410
12	38.2	15.9	12763	6	AX345204
13	38.2	15.9	103507	5	AL928618
14	38.2	15.9	110000	2	AY657029 ¹
15	38.2	15.9	150663	2	AC074258
16	38.2	15.9	184431	10	AC130822
17	38.2	15.9	211697	10	AC122296
18	38	15.8	170559	2	AC026989
19	38	15.8	182997	9	AL355332 Human DNA

c 20	37.8	15.8	97621	9	AL607089	AL607089 Human DNA
21	37.8	15.8	349980	6	AX344560	AX344560 Sequence
c 22	37.6	15.7	9238	6	AX323752	AX323752 Sequence
c 23	37.6	15.7	121922	2	AP000576	AP000576 Homo sapi
24	37.6	15.7	233181	2	AC111558	AC111558 Rattus no
c 25	37.4	15.6	180137	2	BX957257	BX957257 Danio rer
c 26	37.2	15.5	281723	3	PFA929359	AL929359 Plasmidiu
c 27	37	15.4	72000	9	AP003386	AP003386 Homo sapi
28	37	15.4	162397	2	AC011098	AC011098 Homo sapi
29	37	15.4	169993	2	AC144884	AC144884 Papio anu
c 30	37	15.4	210223	10	AC122183	AC122183 Mus muscu
c 31	37	15.4	210567	2	AC074215	AC074215 Homo sapi
32	37	15.4	216031	9	AC011328	AC011328 Homo sapi
33	37	15.4	234203	2	AC121311	AC121311 Mus muscu
34	36.8	15.3	59623	9	AC020591	AC020591 Homo sapi
35	36.8	15.3	110000	2	BX005127 ²	Continuation (3 of
c 36	36.8	15.3	110000	2	BX005127 ³	Continuation (4 of
37	36.8	15.3	177070	2	AC018875	AC018875 Homo sapi
38	36.8	15.3	201823	9	HSG120K12	AL109865 Human DNA
39	36.8	15.3	204483	10	AL671990	AL671990 Mouse DNA
c 40	36.6	15.2	18683	6	AX281292	AX281292 Sequence
c 41	36.6	15.2	18683	6	AX345215	AX345215 Sequence
c 42	36.4	15.2	4895	14	AF156933	AF156933 Hyposoter
c 43	36.4	15.2	54814	4	AC104718	AC104718 Sus scrof
c 44	36.4	15.2	107893	5	BX276180	BX276180 Zebrafish
45	36.4	15.2	146776	2	BX897667	BX897667 Danio rer

ALIGNMENTS

RESULT 1
BD175189
LOCUS BD175189 DNA fragment of a mechiru nutrition nature picha pathotorisu yeast
DEFINITION ICL gene.
ACCESSION BD175189
VERSION BD175189.1 GI:29120883
KEYWORDS JP 2002253235-A/2.
SOURCE Pichia pastoris
ORGANISM Pichia pastoris
REFERENCE Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetaceae; Pichia.
AUTHORS Dias,J.M., Prado,I.V. and Leon,N.C.
TITLE DNA fragment of a mechiru nutrition nature picha pathotorisu yeast
JOURNAL ICL gene
COMMENT Patent: JP 2002253235-A 2 10-SEP-2002;
CENTRO DE INGENIERIA GENETICA Y BIOTECNOLOGIA
OS Pichia pastoris (yeast)
PN JP 2002253235-A/2
PD 10-SEP-2002
PF 28-MAY-2000 JP 2001159321
PI JAVIER MENENDEZ DIAS,IRIS VALDEZ PRADO,NELSON CABRERA LEON PC
C12N15/09,C12N15/00
CC Secuencia que contiene la region 3' no codificante del gen CC
ICL.
FH Key Location/Qualifiers
FT terminator Location/Qualifiers
1..360
/organism="Pichia pastoris"
/mol_type="genomic DNA"
/db_xref="taxon:4922"

Query Match 100.0%; Score 240; DB 6; Length 360;
Best Local Similarity 100.0%; Pred. No. 2.4e-52;
Matches 240; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTAATAGGAGTTCCTAAGTAGTTAAGATAATTCAGCTTGAGGTATTATAGATTGTGTGT 60
|||||

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 13, 2004, 14:02:37 ; Search time 226 Seconds
(without alignments)
5574.606 Million cell updates/sec

Title: US-09-864-486A-2
Perfect score: 240
Sequence: 1 gtaataggagttcctaagta.....aaaaaaaaaacttaacatcaact 240

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq 23Sep04:*

- 1: geneseqm1980s:*
- 2: geneseqm1990s:*
- 3: geneseqm2000s:*
- 4: geneseqm2001as:*
- 5: geneseqm2001bs:*
- 6: geneseqm2002as:*
- 7: geneseqm2002bs:*
- 8: geneseqm2003as:*
- 9: geneseqm2003bs:*
- 10: geneseqm2003cs:*
- 11: geneseqm2003ds:*
- 12: geneseqm2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	238.4	99.3	360	ABK13231	Abk13231 P. pastor
2	38.2	15.9	12763	ABL32302	ABL32302 Human imm
3	37.6	15.7	9238	ABK28366	ABK28366 DNA trans
4	37	15.4	5155	ABN87364	ABN87364 Human lip
5	36.8	15.3	32186	AAS34422	AAS34422 Human DNA
6	36.8	15.3	38855	ADN41808	Adn41808 Novel hum
7	36.6	15.2	18683	ABL32313	ABL32313 Human imm
8	36.6	15.2	18683	ABL32313	ABL32313 Human imm
9	36	15.0	6127	ABL54334	ABL54334 Chemicall
10	36	15.0	6127	ABL34449	ABL34449 Human met
11	36	15.0	9919	ABL70120	ABL70120 Chemicall
12	36	15.0	9919	AAV04631	AAV04631 Arabidops
13	35.8	14.9	6123	ABL32821	ABL32821 Human imm
14	35.8	14.9	6509	ABL32226	ABL32226 Human imm
15	35.2	14.7	7276	ABL32903	ABL32903 Human imm
16	35	14.6	4570	ABL29394	ABL29394 Drosophil
17	35	14.6	6987	ABL29562	ABL29562 Drosophil
18	34.8	14.5	5814	ABL32974	ABL32974 Human imm
19	34.6	14.4	16688	ABL32320	ABL32320 Human imm
20	34.2	14.3	5942	ABZ10233	ABZ10233 Haematopo
21	34.2	14.3	5942	ABZ10147	ABZ10147 Haematopo

C	22	34.2	14.3	5942	10	ADE84143	Ade84143 Human lym
	23	34.2	14.3	15161	6	ABL70458	ABL70458 Chemicall
	24	34.2	14.3	15161	6	AAS61423	Aas61423 Human gen
C	25	34.2	14.3	15872	4	AAS46520	Aas46520 Tumour su
	26	34	14.2	1900	3	AAF15888	Aaf15888 Human pro
C	27	34	14.2	6360	6	ABL34398	ABL34398 Human imm
	28	33.8	14.1	606	5	ABV49913	Abv49913 Human pro
	29	33.8	14.1	7833	10	ADB54184	Adb54184 Pretreate
C	30	33.8	14.1	7833	10	ADB54312	Adb54312 Pretreate
	31	33.8	14.1	7833	10	ADE37769	Ade37769 Human che
C	32	33.8	14.1	7833	10	ADE37779	Ade37779 Human che
C	33	33.8	14.1	8718	6	ABL33273	ABL33273 Human imm
	34	33.6	14.0	799	4	AH31386	Aah31386 Human sec
	35	33.6	14.0	1325	6	ABZ21184	Abz21184 Human BCK
C	36	33.6	14.0	6196	6	ABL33864	ABL33864 Human imm
C	37	33.6	14.0	16766	6	ABL34157	ABL34157 Human imm
	38	33.6	14.0	110000	5	AAF84800	Aaf84800 Nuclcotid
C	39	33.4	13.9	6963	6	ABL32978	ABL32978 Human imm
C	40	33.4	13.9	14798	6	ABL33032	ABL33032 Human imm
C	41	33.4	13.9	15121	6	ABN80239	ABn80239 Human che
C	42	33.4	13.9	17137	6	ABL32191	ABL32191 Human imm
C	43	33.2	13.8	6365	6	ABK31281	ABk31281 Signal tr
	44	33.2	13.8	9110	4	AAS46394	Aas46394 Tumour su
C	45	33.2	13.8	110000	3	AAF22305_01	Continuation (2 of

ALIGNMENTS

RESULT 1
ABK13231
ID ABK13231 standard; DNA; 360 BP.
XX

AC ABK13231;

DT 23-APR-2002 (first entry)

DE P. pastoris isocitrate lyase gene 3' UTR.

KW Isocitrate lyase; ICL; 3' UTR; ds; yeast; heterologous gene expression.

XX Pichia pastoris.

PN EP1162266-A2.

PD 12-DEC-2001.

PF 28-MAY-2001; 2001EP-00202009.

PR 26-MAY-2000; 2000CU-00000122.

XX (INGG-) CENT ING GENETICA & BIOTECNOLOGIA.

PI Menendez Diaz J, Valdes Prado I, Cabrera Leon N;

XX WPI; 2002-084415/12.

PT New recombinant DNA fragment for protein production, contains an isocitrate lyase encoding gene from Pichia pastoris, which regulates heterologous gene expression in yeasts, when operably linked to the fragment.

XX Claim 1; Page 8; 18pp; English.

CC The invention relates to a recombinant DNA fragment (I) comprising a sequence (S) of 684bp (the 5' region of the Pichia pastoris isocitrate lyase, ICL, gene including the promoter) or 360bp (the 3' UTR including a terminator), given in the specification, where (I) belongs to an isocitrate lyase encoding gene (ICL) from Pichia pastoris, which is able to regulate heterologous gene expression in yeasts, when it is operably linked to (I), and including regulatory elements necessary for heterologous gene expression. The DNA fragment is useful for expressing heterologous proteins in Pichia pastoris, for expressing a foreign gene,

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 13, 2004, 14:11:07 ; Search time 45.2 Seconds
(without alignments)
3774.100 Million cell updates/sec

Title: US-09-864-486A-2
Perfect score: 240
Sequence: 1 gtaataggagttcctaagta.....aaaaaaaaacttaacatcact 240

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 35539441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/prodata/1/ina/5A_COMB.seq: *
2: /cgn2_6/prodata/1/ina/5B_COMB.seq: *
3: /cgn2_6/prodata/1/ina/6A_COMB.seq: *
4: /cgn2_6/prodata/1/ina/6B_COMB.seq: *
5: /cgn2_6/prodata/1/ina/6C_COMB.seq: *
6: /cgn2_6/prodata/1/ina/6D_COMB.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	36	15.0	9919	3	US-08-880-179-1
C 2	33	13.8	6583	4	US-10-204-708-25
C 3	32.6	13.6	65933	4	US-09-544-3988-11
C 4	32.6	13.6	65933	4	US-09-543-771-11
C 5	32	13.3	11049	4	US-10-204-708-22
C 6	31.6	13.2	6306	4	US-10-204-708-50
C 7	31.4	13.1	1347	3	US-09-134-001C-218
C 8	31.4	13.1	1347	4	US-09-710-279-1727
C 9	31.4	13.1	4069	4	US-09-710-279-3976
C 10	31.2	13.0	164	4	US-09-621-976-16692
C 11	31.2	13.0	669	4	US-09-799-451-750
C 12	31.2	13.0	2169	3	US-08-981-803-14
C 13	31.2	13.0	2169	3	US-08-981-803-28
C 14	31.2	13.0	2169	3	US-08-983-440-14
C 15	31.2	13.0	2169	3	US-08-983-440-28
C 16	31.2	13.0	2169	4	US-09-367-895-14
C 17	31.2	13.0	2169	4	US-09-367-895-28
C 18	31.2	13.0	5032	4	US-09-344-510B-8
C 19	31.2	13.0	11469	4	US-09-367-895-29
C 20	31.2	13.0	11478	3	US-08-981-803-23
C 21	31.2	13.0	11478	3	US-08-983-440-29
C 22	31	12.9	4351	4	US-09-620-312D-347
C 23	30.8	12.8	458	4	US-09-270-767-3713
C 24	30.8	12.8	458	4	US-09-270-767-18995
C 25	30.8	12.8	791	3	US-08-998-416-346
C 26	30.8	12.8	6070	4	US-10-204-708-10
C 27	30.6	12.8	1026	4	US-09-655-908-19

C 28	30.6	12.8	5501	4	US-10-204-708-37	Sequence 37, Appl
C 29	30.6	12.8	11049	4	US-10-204-708-21	Sequence 21, Appl
C 30	30.6	12.8	11049	4	US-10-204-708-21	Sequence 21, Appl
C 31	30.4	12.7	516	4	US-08-956-171E-597	Sequence 597, Appl
C 32	30.4	12.7	516	4	US-08-781-986A-597	Sequence 597, Appl
C 33	30.4	12.7	11050	4	US-10-204-708-86	Sequence 86, Appl
C 34	30.4	12.7	11811	3	US-09-078-294-7	Sequence 7, Appl
C 35	30.4	12.7	19513	4	US-10-204-708-39	Sequence 39, Appl
C 36	30.2	12.6	972	1	US-07-915-934-1	Sequence 1, Appl
C 37	30.2	12.6	972	1	US-08-325-743-1	Sequence 1, Appl
C 38	30.2	12.6	1674	4	US-09-248-796A-10284	Sequence 10284, A
C 39	30.2	12.6	1722	4	US-09-248-796A-6457	Sequence 6457, A
C 40	30.2	12.6	5666	4	US-10-204-708-29	Sequence 29, Appl
C 41	30.2	12.6	6866	4	US-10-204-708-20	Sequence 20, Appl
C 42	30.2	12.6	193303	4	US-09-497-855A-37	Sequence 37, Appl
C 43	30.2	12.6	193303	4	US-09-497-855A-44	Sequence 44, Appl
C 44	30	12.5	8578	4	US-09-784-358-17	Sequence 17, Appl
C 45	29.8	12.4	516	4	US-09-601-198-79	Sequence 79, Appl

ALIGNMENTS

RESULT 1
US-08-880-179-1/c
; Sequence 1, Application US/08880179
; Patent No. 6091004
; GENERAL INFORMATION:
; APPLICANT: Ryals, John
; APPLICANT: Delaney, Terry
; APPLICANT: Friedrich, Leslie
; APPLICANT: Weymann, Kristianna
; APPLICANT: Lawton, Kay
; APPLICANT: Ellis, Daniel
; APPLICANT: Uknes, Scott
; APPLICANT: Jesse, Taco
; APPLICANT: Vos, Pieter
; TITLE OF INVENTION: GENE ENCODING A PROTEIN INVOLVED IN THE
; TITLE OF INVENTION: SIGNAL TRANSDUCTION CASCADE LEADING TO SYSTEMIC ACQUIRED RES
; TITLE OF INVENTION: IN PLANTS
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6091004artis Corporation
; STREET: 520 White Plains Road, P.O. Box 2005
; CITY: Tarrytown
; STATE: New York
; COUNTRY: USA
; ZIP: 10591
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/880,179
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Weigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: CGC 1909
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919) 541-8587
; TELEFAX: (919) 541-8689
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9919 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 13, 2004, 15:27:22 ; Search time 236 Seconds

(without alignments)
5606.970 Million cell updates/sec

Title: US-09-864-486A-2

Perfect score: 240

Sequence: 1 gtaataggagtcttaagta.....aaaaaaacttaacatcact 240

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4085731 seqs, 2756760397 residues

Total number of hits satisfying chosen parameters: 8171462

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
- 20: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
- 21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	240	100.0	240	11	US-09-864-486A-2
2	240	100.0	360	11	US-09-864-486A-1
c 3	38.2	15.9	12763	15	US-10-311-455-275
4	37.8	15.8	3673778	15	US-10-312-841-1
c 5	37.6	15.7	9238	15	US-10-240-453-240
6	37	15.4	55155	9	US-09-735-933-3
c 7	36.8	15.3	38855	11	US-09-973-278-936
c 8	36.2	15.2	18683	15	US-10-311-455-286
c 9	36.6	15.2	18683	15	US-10-240-452-34
c 10	36	15.0	6127	15	US-10-240-485-2
c 11	36	15.0	9919	13	US-10-079-035-1
12	35.8	14.9	2829	18	US-10-425-115-172689

c 13	35.8	14.9	6123	15	US-10-311-455-794	Sequence 794, App
14	35.8	14.9	6509	15	US-10-311-455-199	Sequence 199, App
c 15	35.4	14.8	3673778	15	US-10-312-841-2	Sequence 2, Appli
c 16	35.2	14.7	7276	15	US-10-311-455-876	Sequence 876, App
c 17	35.2	14.7	3673778	15	US-10-312-841-1	Sequence 1, Appli
18	34.8	14.5	637	13	US-10-027-632-201748	Sequence 201748,
19	34.8	14.5	637	13	US-10-027-632-201749	Sequence 201749,
20	34.8	14.5	637	13	US-10-027-632-201748	Sequence 201748,
21	34.8	14.5	637	15	US-10-027-632-201749	Sequence 201749,
22	34.8	14.5	5814	15	US-10-311-455-947	Sequence 947, App
c 23	34.6	14.4	16688	15	US-10-311-455-293	Sequence 293, App
c 24	34.6	14.4	16688	15	US-10-311-455-293	Sequence 293, App
c 25	34.4	14.3	590	13	US-10-027-632-227310	Sequence 227310,
c 26	34.4	14.3	590	13	US-10-027-632-227311	Sequence 227311,
c 27	34.4	14.3	590	13	US-10-027-632-227312	Sequence 227312,
c 28	34.4	14.3	590	15	US-10-027-632-227310	Sequence 227310,
c 29	34.4	14.3	590	15	US-10-027-632-227311	Sequence 227311,
c 30	34.2	14.3	590	15	US-10-027-632-227312	Sequence 227312,
c 31	34.2	14.3	5942	18	US-10-425-115-74299	Sequence 74299, A
c 32	34.2	14.3	5942	18	US-10-473-126-373	Sequence 287, App
c 33	34.2	14.3	15161	16	US-10-473-126-373	Sequence 373, App
c 34	34.2	14.3	15872	16	US-10-221-714A-242	Sequence 386, App
c 35	34	14.2	1900	9	US-09-925-300-123	Sequence 442, App
c 36	34	14.2	6960	15	US-10-311-455-2371	Sequence 123, App
c 37	33.8	14.1	338	16	US-10-242-535A-13681	Sequence 2371, Ap
c 38	33.8	14.1	338	16	US-10-085-783A-13681	Sequence 13681, A
c 39	33.8	14.1	8718	15	US-10-311-455-1246	Sequence 1246, Ap
c 40	33.6	14.0	6196	15	US-10-311-455-1837	Sequence 1837, Ap
c 41	33.6	14.0	16766	15	US-10-311-455-2130	Sequence 2130, Ap
c 42	33.4	13.9	6963	15	US-10-311-455-951	Sequence 951, App
c 43	33.4	13.9	14798	15	US-10-311-455-1005	Sequence 1005, Ap
c 44	33.4	13.9	17137	15	US-10-311-455-164	Sequence 164, App
c 45	33.2	13.8	9110	16	US-10-221-714A-116	Sequence 116, App

ALIGNMENTS

RESULT 1

US-09-864-486A-2
; Sequence 2, Application US/09864486A
; Publication No. US20040106783A1
; GENERAL INFORMATION:
; APPLICANT: Menendez Diaz, Javier
; APPLICANT: Valdes Prado, Chris
; APPLICANT: Cabrera Leon, Nelson
; TITLE OF INVENTION: DNA Fragments of the Methylotrophic Pichia Pastoris Yeast iCL g
; FILE REFERENCE: 976-9
; CURRENT APPLICATION NUMBER: US/09/864,486A
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 240
; TYPE: DNA
; ORGANISM: Pichia pastoris
; US-09-864-486A-2

Query Match 100.0%; Score 240; DB 11; Length 240;
Best Local Similarity 100.0%; Pred. No. 1.9e-54;
Matches 240; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	GTAATAGAGTTCCTAAGTAGTAAAGATAATGACCTTGAGGTATTTATAGATTGTGTGT	60
Db	1	GTAATAGAGTTCCTAAGTAGTAAAGATAATGACCTTGAGGTATTTATAGATTGTGTGT	60
Qy	61	AGGTAAATCATCTGCTCCATCTTACCTTGGTGGGTGACGGGGCGGTGAATAATC	120
Db	61	AGGTAAATCATCTGCTCCATCTTACCTTGGTGGGTGACGGGGCGGTGAATAATC	120
Qy	121	AGTTGCCATGAAGACATTTTACACCTTGTCCACAGAGGTGCGGTCTACTGATTACTACA	180
Db	121	AGTTGCCATGAAGACATTTTACACCTTGTCCACAGAGGTGCGGTCTACTGATTACTACA	180

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 13, 2004, 14:10:32 ; Search time 1511.6 Seconds
(without alignments)
5785.615 Million cell updates/sec

Title: US-09-864-486A-2
Perfect score: 240
Sequence: 1 gcaatagaggtctctaagta.....aaaaaaacttaacatcact 240

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gssi:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	46	19.2	476	1	AI399548 NCSPE6C2T7
C 2	39.8	16.6	329	6	CF329701 NACL--05-
C 3	38.4	16.0	519	5	BU533351 AGENCOURT
C 4	38.4	16.0	568	7	CC847190 ND.L.9714.
C 5	38.2	15.9	395	7	CO819665 CSECS157H
C 6	38.2	15.9	531	7	CF479344 RTW3_23_
C 7	38.2	15.9	757	9	AG495046 Mus muscu
C 8	38	15.8	282	5	BU564984 AGENCOURT
C 9	38	15.8	500	6	CA396793 CS82C02.Y
C 10	37.8	15.8	865	9	CL130913 ISB1-98N2
C 11	37.4	15.6	1201	9	AL106036 Drosophil
C 12	37.2	15.5	168	4	BI744958 rk94d10.Y
C 13	37.2	15.5	265	4	BM343692 rr41g09.Y
C 14	37.2	15.5	1101	9	AL063921 Drosophil
C 15	37	15.4	127	1	AI865961 wk87g01.X
C 16	37	15.4	563	7	CN985874 C1448.125
C 17	37	15.4	707	6	CF207352 CAB20001
C 18	37	15.4	870	5	BQ933590 AGENCOURT
C 19	36.8	15.3	646	5	BU485437 603468625
C 20	36.8	15.3	686	5	BU208472 604151831
C 21	36.8	15.3	733	7	CR577504 CR577504
C 22	36.8	15.3	886	9	CL134229 ISB1-104K
C 23	36.8	15.3	938	7	CK020362 AGENCOURT
C 24	36.8	15.3	1204	9	AL106628 Drosophil

25	36.8	15.3	10560	8	AQ839853
26	36.6	15.2	296	5	BQ106987
27	36.6	15.2	511	9	TA128B12Q
28	36.6	15.2	703	9	CC503344 CH240_342
29	36.6	15.2	920	9	CNS0046U
30	36.6	15.2	1101	9	CNS0024Y
31	36.4	15.2	615	2	BF299319
32	36.4	15.2	662	9	EX225916
33	36.4	15.2	1101	9	CNS00296
34	36.4	15.2	1101	9	CNS016TA
35	36.2	15.1	270	6	CB218381
36	36.2	15.1	286	7	H23747
37	36.2	15.1	537	5	BX522361
38	36.2	15.1	585	9	CE699052
39	36.2	15.1	687	1	AL855043
40	36.2	15.1	722	4	BM690925
41	36.2	15.1	783	3	AY067688
42	36	15.0	605	2	AW307435
43	36	15.0	613	9	BX213573
44	35.8	14.9	481	8	BZ618173
45	35.8	14.9	486	1	AL714653

ALIGNMENTS

RESULT 1
AI399548/c
LOCUS NCSPE6C2T7 476 bp mRNA linear EST 08-FEB-1999
DEFINITION NCSPE6C2T7 Subtracted Perithecial Neurospora crassa cDNA clone SP6C2
3' similar to ORF YLR251W, probable membrane protein, mRNA
sequence.
ACCESSION AI399548
VERSION AI399548.1 GI:4242635
KEYWORDS EST.
SOURCE Neurospora crassa
ORGANISM Neurospora crassa
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
REFERENCE 1 (bases 1 to 476)
AUTHORS Nelson,M.A., Kang,S., Braun,E.L., Crawford,M.E., Dolan,P.L., Leonard,P.M., Mitchell,J., Armiijo,A.M., Bean,L., Blueves,E., Cushing,T., Errett,A., Fleharty,M., Gorman,M., Judson,K., Miller,R., Ortega,J., Pavlova,I., Perea,J., Todisco,S., Trujillo,R., Valentine,J., Wells,A., Werner-Washburne,M., Yazzie,S. and Natvig,D.O.
Expressed sequences from conidial, mycelial, and sexual stages of Neurospora crassa
Fungal Genet. Biol. 21, 348-363 (1997)
97435349
9290248
Contact: Natvig,D.O./Nelson,M.A.
Department of Biology
University of New Mexico
Castetter Hall, Albuquerque, NM 87131, USA
Tel: 505 277 3411
Fax: 505 277 0304
Email: ngp@biology.unm.edu.
Location/Qualifiers
1..476
/organism="Neurospora crassa"
/mol_type="mRNA"
/strain="fl a (FGSC 4347)"
/db_xref="taxon:5141"
/clone="SP6C2"
/sex="Mating type a (fluffy), fertilized"
/tissue_type="Perithecia"
/dev_stage="Fruiting Body"
/lab_host="E. coli"
/clone_lib="Subtracted Perithecial"
/note="Vector: pBluescript SK (-); Site 1: EcoRI; Site 2: XhoI; mRNA isolated from 5 day old perithecia (fruiting bodies) of the fluffy strain fl a (Mating type a),